Analysis of QTL Studies Related to Yield and Vigour Traits Carried out With Different Cocoa Genotypes

Didier Clément[1], Ange-Marie Risterucci[2] and Claire Lanaud[2]

1 - Cirad-Cp, TA 80/02, Avenue Agropolis, 34398 Montpellier Cedex 5, France
2 - Cirad-Amis, TA 40/03, Avenue Agropolis, 34398 Montpellier Cedex 5, France

Abstract

Theobroma cacao is mainly cultivated by small growers, especially in the main production area in West Africa. The sustainability of cocoa cultivation will be improved if farmers have access to new planting material with improved agronomic traits such as yield, vigour, pest, and disease resistance. Progress in breeding programmes to accumulate favourable alleles for these traits can be accelerated using molecular marker techniques which allow more direct access to the genome. The technology has developed rapidly and over the last ten years studies have been carried out to map QTLs for agronomic traits in several plant species. In cocoa, a number of progenies have been mapped and several QTL related to resistance to *Phytophthora* spp. and to yield components have been detected. The comparison of the different linkage maps of cocoa is possible through specific markers (RFLP and microsatellites) mapped on to a reference map containing 473 markers. The purpose of this paper is to analyse available results on the detection of QTL for yield and vigour traits and the co-location of the QTL identified in different parental genotypes. Methodological approaches to the detection of QTL are also presented. Perspectives for further research on mapping of yield and vigour traits and the possible use of molecular markers in the selection for these traits in cocoa are discussed.

Introduction

Most cocoa is produced by small holders who cultivate the trees in addition to food crops for family consumption. The income they gain from cocoa is often saved to support them in later life. Africa supplies nearly 70% of the world's production of cocoa with Côte d'Ivoire as the principal producer. Improved planting material, better adapted to the local environment, and disease conditions, can be produced through breeding programmes designed to accumulate favourable alleles for the main agronomic traits. In cocoa, the selection cycle normally requires 6 to 8 years of observations. Therefore the development of effective early screening methods to accelerate progress is of great interest. Cocoa breeders are also investigating the use of successive cycles of selection, for example a recurrent selection program has been initiated in Côte d'Ivoire to improve lower Amazon Forastero and upper Amazon Forastero populations by successive cycles of selection inside each population (Clément *et al.* 1994; Eskes *et al.* 1995).

Recent progress in molecular marker technology now allows more direct access to the genotype and the use of molecular markers in the genetic mapping of the main agronomic traits has been considerably developed in recent years. A new field of research in cocoa breeding has opened. The first cocoa linkage map was presented by Lanaud *et al.* (1995) from a progeny resulting from the Côte d'Ivoire cocoa breeding programme and a high density linkage map was established recently by Risterucci *et al.* (2000). Genetic mapping of agronomic traits (disease resistance and yield components) was carried out with this reference progeny by Lanaud *et al.* (1999) and with other progenies by Flament *et al.* (2000), by Crouzillat *et al.* (1996, 2000) and by Clément *et al.* (2000).

In the present paper we report the main results of studies to detect QTL related to yield components, to attack by black pod disease and to vigour traits. Furthermore,

comparisons with results from other studies are made and perspectives for the use of molecular markers in cocoa selection programmes discussed.

Material and methods

Plant material

The progenies used for the QTL studies compared in this paper and the references of the publications related to these studies are presented in Table 1. Pound 12, IMC 78, T 60/887, and UPA 402 are Upper Amazon Forastero genotypes. Pound 12 and IMC 78 are genotypes collected by Pound (1938, 1943), UPA 402 is a result of sib-crossing between two genotypes of the T 87 progeny (NA 34 x IMC 60), and T 60/887 is a progeny of PA 7 x NA 32. UF 676, DR 1, and S 52 are Trinitario clones. IFC 1 and IFC 5 are Lower Amazon Forastero selections from Côte d'Ivoire; IFC 5 has probably received some introgression of genes from African Trinitario. Catongo is a highly homozygous Lower Amazon Forastero selected in the Bahia State in Brazil. The number of trees observed in each progeny varied between 55 and 181.

Table 1. Cocoa progenies used in QTL studies for agronomic traits

Progenies	Countries	Number of trees	References
UPA402 x UF676	Côte d'Ivoire	181	Lanaud *et al.* (1995) Risterucci *et al.* (2000)
Catongo x Pound12	Costa Rica	55	Crouzillat *et al.* (1999)
T60/887 x (IFC5 IFC1)	Côte d'Ivoire	112	Flament *et al.* (2000)
DR1 x Catongo	Côte d'Ivoire	107	Clément *et al.* (2000)
S52 x Catongo	Côte d'Ivoire	101	Clément *et al.* (2000)
IMC78 x Catongo	Côte d'Ivoire	128	Clément *et al.* (2000)

Quantitative traits

Yield and vigour data were obtained over several years during the juvenile production phase (first seven years of production), during the mature phase or during both phases (Table 2). The following traits were analysed in the different studies:

- Yield (mean wet bean weight per tree);
- The average weight of one pod or the pod index (these two traits are similar);
- Vigour observed on all adult trees (trunk diameter-TD-, trunk circumference-TC-, or canopy width-CW). Early vigour was measured only for the DR 1, S 52, and IMC 78 progenies two years after the planting (stem diameter-SD).

Mapping analysis and QTL detection

A total of 473 markers were used for the genetic mapping of UPA 402 x UF 676, currently the most saturated map of the cocoa genome. RFLP probes and microsatellites were used as co-dominant markers and AFLP and RAPD as dominant markers. Genetic maps built from progenies with heterozygous parents (double pseudo test cross) such as UPA 402 x UF 676, were made generally using Joinmap version 1.4 (Stam 1993). A genetic map of Catongo x Pound 12 (Crouzillat *et al.* 1996) was established with MAPMAKER (Lander *et al.* 1987). For all these maps, the Kosambi mapping function was used.

RFLP and microsatellite markers allowed linkage groups to be identified with individual chromosomes in UPA 402 x UF 676 (Lanaud *et al.* 1995). However the genetic map of Catongo x Pound 12, established by Crouzillat *et al.* (1996), uses a different chromosome numbering system to that established by Lanaud *et al.* (1995).

Exchanges of RFLP and microsatellite markers has allowed the correspondence between the two chromosome numbering systems to be established and this correspondence is shown in Table 3. In this present study, the UPA 402 x UF 676 map was used as consensus map for locating the QTL identified in the different parental genotypes.

Table 2. Yield and vigour traits observed in the different studies

		Juvenile phase	Mature phase
Yield	**UPA402 x UF676**	4 years	-
	Pound12	7 years	8 years
	T60/887	2 years	
	DR1	-	9 years
	S52	-	9 years
	IMC78	-	9 years
Vigour	**UPA402 x UF676**	-	TC
	Pound12	-	TC and CW
	T60/887	-	TC and CW
	DR1	SD	TC and CW
	S52	SD	TC and CW
	IMC78	SD	

SD: stem diameter
TC: trunk circumference
CW: canopy width

QTL mapping was carried out using the Simple Interval Mapping (SIM) technique proposed by Lander and Botstein (1989). Composite Interval Mapping (CIM), developed by Zeng (1994), was applied only for QTL identified in the DR 1, S 52, and IMC 78 parents (Clément *et al.* 2000). A LOD threshold value of 2 was generally applied to declare that a specific QTL was significant. In the study of Clément *et al.* (2000) the LOD threshold values were fixed by the Churchill and Doerge method (1994).

Table 3. Correspondences between UPA402 x UF676 and Catongo x Pound12 maps

UPA402 x UF676 (1)	Catongo x P12 (2)
Chromosome 1	Chromosome 3
Chromosome 2	Chromosome 4
Chromosome 3	Chromosome 6
Chromosome 4	Chromosome 5
Chromosome 5	Chromosome 9
Chromosome 6	Chromosome 7
Chromosome 7	Chromosome 1
Chromosome 8	Chromosome 8
Chromosome 9	Chromosome 2
Chromosome 10	Chromosome 10

(1) Lanaud *et al.* 1995
(2) Crouzillat *et al.* 1996

Comparison of results

Phenotypic correlation

A strong significant correlation between yield and vigour for adult trees was found in the different studies (Table 4). This is to be expected since competition will have had its effect by the time the trees reach this adult stage. Except for the S 52 x Catongo progeny, the correlation between yield and vigour at a young age was not significant. In general, early vigour is more significantly correlated with the first years of production than with production in later years (Eskes *et al.* 1995).

Table 4. Phenotypic correlation between yield and vigour traits

Parents		Vigour		
		Stem diameter	Trunk circumference	Canopy width
DR1	Yield	NS	0.71***	0.47**
S52		0.38**	0.79**	0.42**
IMC78		NS	0.66**	0.48**
Pound12		(no data)	0.56***	(no data)

Pearson coefficient correlation: significant at $p<0.01$ ** and $p<0.001$***

Genetic maps

For crosses with Catongo (highly homozygous), the markers reflect only the heterozygosity of the parents: Pound 12, DR 1, S 52, and IMC 78. This is also broadly the case with T 60/887, IFC 1, and IFC 5 which are all fairly homozygous. For the UPA 402 x UF 676 map, the marker segregation mainly reflects the high level of heterozygosity of UF 676. The characteristics of each of the maps are shown in Table 5.

Table 5. Genetic maps characteristics

	Number of markers	LG	Map length in cM	Average distance between 2 markers
UPA402 x UF676	473	10	887	1.9
Pound12	158	10	772	4.8
T60/887	184	11	793	4.3
DR1	192	9	653	3.4
S52	138	11	589	4.8
IMC78 x	223	10	721	3.2

QTL detection

SIM and CIM analyses were carried out for QTL detection with DR 1, S 52, and IMC 78 (Clément *et al.* 2000). In the other studies, QTL analyses were carried out using SIM only. The comparison of results obtained using SIM and CIM showed differences in power (LOD score of the peak) and efficiency (estimation of the R^2). In some cases the percentage of the phenotypic variation explained by the QTL (R^2 values) estimated in CIM analyses were higher than those obtained using SIM, but lower in others (Clément *et al.* 2000).

QTL related to yield, average weight of one pod and vigour traits

The most significant QTL for yield, average weight of one pod and vigour traits were mainly detected on chromosomes 1, 4, and 5 in the different heterozygous parents.

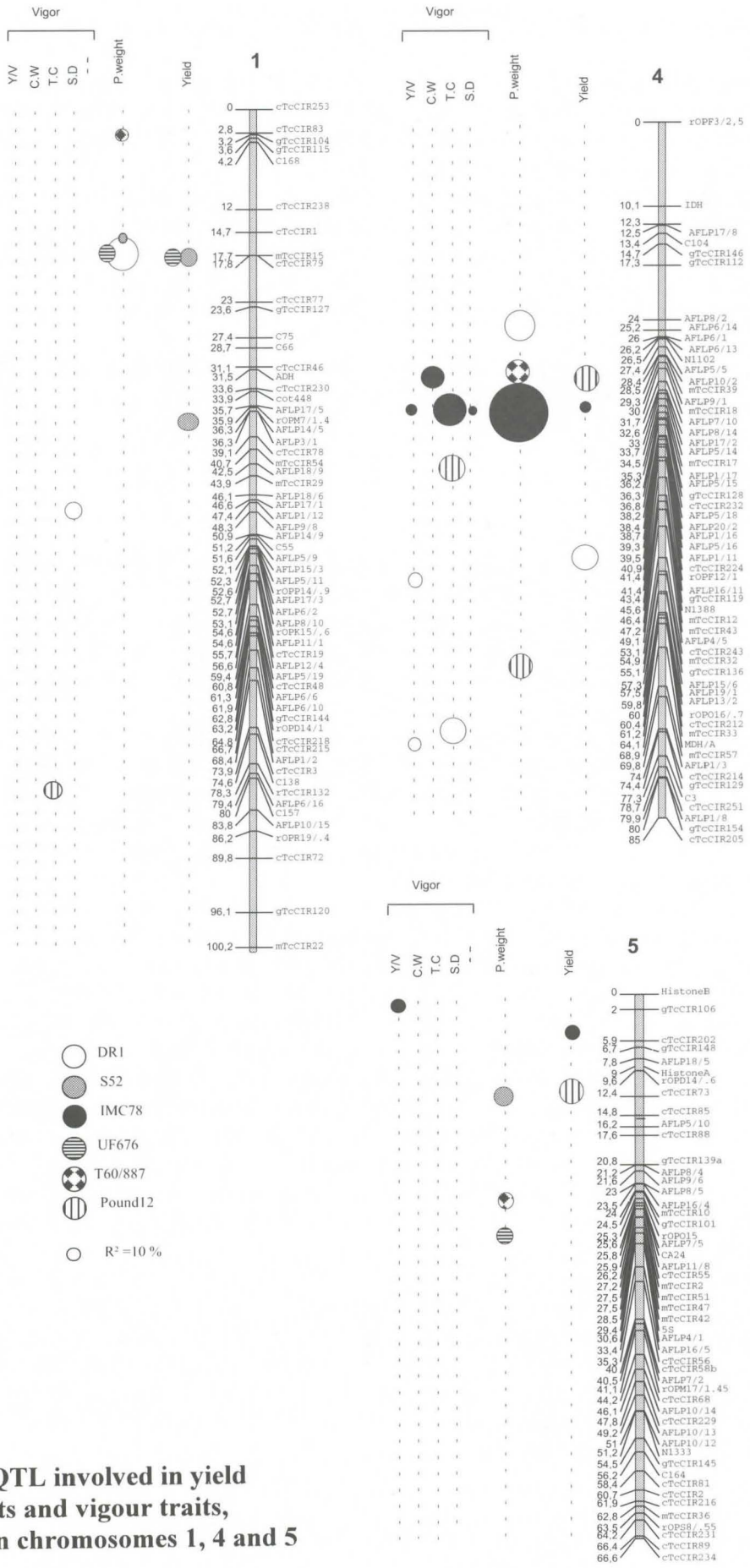

Figure 1. QTL involved in yield components and vigour traits, detected on chromosomes 1, 4 and 5

The results are illustrated in Figure 1, based on the consensus map of UPA 402 x UF 676.

QTL related to yield were detected in the same region of chromosome 1 (around mTcCIR15) for two Trinitario genotypes (UF 676 and S 52). QTL related to yield for two Upper Amazon Forastero genotypes (POUND 12 and IMC 78) were also identified in the same regions, but of chromosomes 4 and 5.

QTL related to the average weight of one pod have been identified in the same region as QTL for yield (Figure 1). In the case of chromosome 4, QTL related to the average weight of one pod were detected in the same region for T 60/887 and IMC 78. A QTL was detected close to this region in DR 1. The QTL located in IMC 78 explained 43.5% of the phenotypic variation; it might therefore involve a major gene (Clément *et al.* 2000).

On chromosome 4; several QTL for yield components and vigour traits were detected in the same region (Figure 1)

Discussion

Methods

Genetic mapping of yield, its components and other agronomic traits must be carried out using data gathered from field trials over several years. Most of the progenies analysed were planted in existing hybrid variety trials. This implies that, in most cases, the statistical analysis has been carried out using data from less than 100 trees. It is considered now that QTL analyses require observations on progenies with at least 200 individuals.

QTL studies based on the analysis of several progenies sharing common parental genotypes require fewer individuals for each progeny. This method is interesting because it allows the stability of the QTL in different genetic backgrounds to be estimated. Simulation studies on poplar showed that better results could be expected using progenies with common parents (*e.g.* factorial mating designs) than using unrelated progenies (Muranty *et al.* 1996).

Different chromosome numbering systems are currently being used in the two genetic maps of cocoa. It is suggested that the numbering system established for the UPA 402 x UF 676 progeny should be universally adopted (Risterucci *et al.* 2000).

Microsatellite markers reveal a higher level of polymorphism than RFLP markers (Risterucci *et al.* 2000; Clément *et al.* 2000). They also require less DNA and some of the analyses can be automated. The microsatellite technology appears to be currently the most appropriate method for genetic mapping and QTL analyses in cocoa. Indeed, microsatellite markers allow easy identification of the linkage groups and comparison of maps from different progenies. They have also been used in genetic diversity studies and to confirm the identity of clones and seedling progeny. Moreover, microsatellite marker technology will be easily transferable to laboratories in tropical countries (since the technique does not involve the use of radioactivity) and applied to Marked Assisted Selection (MAS).

The Simple Interval Mapping (SIM) technique proposed by Lander and Botstein (1989) has been the detection method generally used for the detection of QTL in cocoa. However, the Composite Interval Mapping (CIM) approach increases detection power and improves the estimation of the phenotypic variation explained by the QTL. With CIM analyses, 5 to 10 markers, given by the forward, and backward regression (co-factors added to the model), are generally used. This method is being increasingly applied in QTL mapping analyses.

QTL identified

Genetic diversity studies on the Trinitario group have shown that this group resulted from hybridisation between almost homozygous Criollo and Forastero individuals

(Motamayor *et al.* 2000). In this situation, we can suppose that a linkage disequilibrium between molecular markers and agronomic traits may have been maintained. Indeed, common QTLs related to yield components have been found in different Trinitario clones (DR 1, S 52, and UF 676). This could mean that there is a good chance that the QTL identified in one Trinitario type will also apply to other Trinitario genotypes.

For the Pound 12 and IMC 78 progenies, yield was observed over several years in Costa Rica and Côte d'Ivoire, respectively. The co-localised QTL for yield of these two Forasteros were the ones which were also the most stable in time. This co-location of QTLs between IMC 78 and POUND 12 could possibly also be due to the similar genetic origin of these genotypes. Indeed, according to Pound (1943), the genotypes collected on the Nanay river (as POUND 12) grew not far from the area where the IMC (Iquitos Mixed Calabacillo) trees were collected. Recent diversity studies on Upper Amazon Forastero genotypes showed that the genetic distance between Nanay (NA) and IMC clones was relatively low (Sounigo *et al.* 2000). Therefore, a genetic linkage disequilibrium may have been maintained between markers and genes involved in yield components between IMC and NA genotypes.

The QTL for the average pod weight with the IMC 78 parent was detected on chromosome 4 and explained 43.5% of the phenotypic variation. Close to it, a QTL related to the same trait was identified in DR 1 (R^2=22). Studies on the genetic control of fruit traits (weight and size) have been carried out on other species such as tomato (Grandillo *et al.* 1999; Ku *et al.* 1999). An example is given by Ku *et al.* (1999) on the genetic control of fruit length and the constriction at the stem end of the fruit. We can suggest a similar situation concerning genes involved in the genetic control of pod form (weight, shape) in cocoa especially for the QTL located in the common region of chromosome 4 of these two types. Research to establish a candidate gene for this yield component could be envisaged.

Perspectives

Molecular markers have opened a new era of more efficient selection for quantitative traits. In this study, QTLs involved in important traits for breeding were identified. The co-localisation of some of them in progenies from parents belonging to the either the same or different genetic groups confirms the stability of some of the QTL identified. This is a favourable situation to consider marker assisted selection (MAS).

MAS allows breeders to follow two main approaches. The first approach consists of monitoring the accumulation of favourable genes in one genotype in back-cross progenies (Marked Assisted Back-Cross). The second allows for better assessment of the genotypic value of individuals from the marker genotype, allowing for Marker Assisted Recurrent Selection (MARS). This second approach seems more appropriate for the use of molecular markers in cocoa breeding programmes. Various applications of MARS were recently developed by Gallais *et al.* (2000).

The advantage of selection based only on markers is that the selection cycle can be considerably shortened. In order to use all sources of genetic variability from marked and unmarked QTLs it seems, however, better to use Marker Assisted Selection, combining molecular score and phenotypic value in an index, such as that defined by Lande and Thompson (1990).

References

Clément D., A.B. Eskes, O. Sounigo and J.A.K. N'Goran. 1994. Amélioration génétique du cacaoyer en Côte d'Ivoire. Présentation d'un nouveau schéma de sélection. Pages 451-455 *in* Proceedings of the11th International Cocoa Research Conference, July 18-24, 1993, Yamoussoukro, Ivory Coast. Cocoa Producer's Alliance, Lagos, Nigeria.

Clément D., A.M. Risterucci, L. Grivet, J-C. Motamayor, J.A.K. N'Goran and C. Lanaud. 2000. Cartographie génétique de caractères liés au rendement, à la vigueur, à la résistance à *P. palmivora* et aux fèves, chez *Theobroma cacao.* In press *in* Proceedings of the 13th International Cocoa Research Conference, October 9-14, Kota Kinabalu, Malaysia.

Churchill G.A. and R.W. Doerge. 1994. Empirical threshold values for Quantitative Trait Mapping. Genetics 8: 963-971.

Crouzillat D., E. Lerceteau, V. Pétiard, J. Morena, H. Rodriguez, D. Walker, W. Phillips, C. Ronning, R. Schnell, J. Osei and P. Fritz. 1996. *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. Theor. Appl. Genet. 93: 205-214.

Crouzillat D., B. Ménard, A. Mora, W. Phillips and V. Pétiard. 2000. Quantitative trait analysis in *Theobroma cacao* L. using molecular markers. Euphytica 114: 13-23.

Eskes A.B., D. Paulin, D. Clément, J.A.K. N'Goran, O. Sounigo, P. Lachenaud, C. Cilas, D. Berry and A. Yagmpam. 1995. Méthodes de sélection et connaissances acquises en génétique par les programmes d'amélioration du cacaoyer en Côte d'Ivoire et au Cameroun. Pages 41-56 *in* Proceedings of the International workshop on Cocoa Breeding Strategies. October 18-19, 1994, Kuala Lumpur, Malaysia. INGENIC, United Kingdom.

Flament M.H., I. Kébé, D. Clément, I. Pierretti, A.M. Risterucci, J.A.K. N'Goran, C. Cilas, D. Despréaux and C. Lanaud. 2000. Genetic mapping of resistance factors to *Phytophthora palmivora* in cocoa. Génome. In press.

Gallais A., A. Charcosset, I. Goldringer, F. Hospital and L. Moreau. 2000. Progress and prospects for marker-assisted selection. Quantitative Genetics and Breeding Methods. *In* Proceedings of the XIth Meeting of the Section Biometrics in Plant Breeding, August 30-September 01 2000, Paris, France. In press

Grandillo S., H.M. Ku and S.D. Tanksley. 1999. Identifying the loci responsible for natural variation in fruit size and shape in tomato. Theor. Appl. Genet. 99: 978-987.

Ku H-M., S. Doganlar, K.Y. Chen and S.D. Tanksley.1999. The genetic basis of pear-shaped tomato fruit. Theor. Appl. Genet. 9: 844-850.

Lanaud C., A.M. Risterucci, J.A.K. N'Goran, D. Clément, M. Flament, V. Laurent and M. Falque. 1995. A genetic linkage map of *Theobroma cacao* L. Theor. Appl. Genet. 91: 987-993.

Lanaud C., I. Kébé I, A.M. Risterucci, D. Clément, J.A.K. N'Goran, L. Grivet, M. Tahi, C. Cilas, I. Pieretti, A.B. Eskes and D. Despréaux. 1999. Mapping quantitative trait loci (QTL) for resistance to *Phytophthora palmivora* in *T. cacao* L. Pages 99-105 *in* Proceedings of the 12th International Cocoa Research Conference, November 17-23, 1996, Salvador, Bahia, Brazil

Lande R. and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124: 743-756.

Lander E., P. Green, J. Abrahamson, A. Barlow, M. Daly, S. Lincoln and L. Newburg. 1987. Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174-181.

Lander E.S. and D. Botstein. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121: 185-199.

Motamayor J.C., A.M. Risterucci and C. Lanaud. 2000. Cacao domestication II: progenitor germplasm of most cultivated cacao until 1950. In press *in* Proceedings of the 13th International Cocoa Research Conference, October 9-14, Kota Kinabalu, Malaysia.

Muranty H. 1996. Power of tests for quantitative trait loci detection using full-sib families in different schemes. Heredity, 76: 156-165.

Pound F.J. 1943. Cocoa and witches' broom disease: report on a recent visit to the Amazon territory of Peru. The archives of cocoa research, volume 1, (Toxopeus, H., ed), ACRI-IOCC, London, United Kingdom.

Risterucci A.M., L. Grivet, J.A.K. N'Goran, I. Pieretti, M-H. Flament and C. Lanaud. 2000. A high-density linkage map of *Theobroma cacao* L. Theor. Appl. Genet. 101: 948-955.

Sounigo O., Y. Christopher, S. Ramdahin, R. Umaharan and A. Sankar. 2000. Evaluation and use of the genetic diversity present in the International Cocoa Genebank, Trinidad (ICG T). *In* Proceedings of the 3rd INGENIC International Workshop on the New Technologies and Cocoa Breeding, October 16-17, Kota Kinabalu, Sabah, Malaysia. In press.

Stam P. 1993. Construction of integrated genetic linkage maps by means by a new computer package: JoinMap. The Plant Journal 3: 739-744.

Zeng Z.B. 1994. Precision mapping of quantitative trait loci. Genetics 136: 1457-1468.